

# Continuity estimates on the Tsallis relative entropy

Alexey E. Rastegin

*Department of Theoretical Physics, Irkutsk State University, Gagarin Bv. 20, Irkutsk 664003, Russia*

Continuity properties of the Tsallis relative entropy are examined. The monotonicity of the quantum  $f$ -divergence leads to a consequence which is ready for estimating this measure from below. For order  $\alpha \in (0; 1)$ , a family of lower continuity bounds of Pinsker type is obtained. For  $\alpha > 1$  and the commutative case, upper continuity bounds on the relative entropy in terms of the minimal probability in its second argument are derived. Both the lower and upper bounds presented are reformulated for the case of Rényi's entropies. The Fano inequality is extended to Tsallis' entropies for all  $\alpha > 0$ . The deduced bounds on the Tsallis conditional entropy are used for obtaining inequalities of Fannes type.

PACS numbers: 03.67.-a, 03.65.Ta

Keywords: Tsallis entropy, Rényi entropy, Pinsker inequality, Fano inequality, Fannes inequality, trace distance, convexity

## I. INTRODUCTION

Physical systems involving long-range interactions, long-time memories, or fractal structures can hardly be treated within the traditional background of statistical physics. The Tsallis entropies have been widely adopted in this direction [16]. As a rule, stationary states of such systems are described by one-parametric extensions of the Zipf-Mandelbrot power-law distribution. Generalized entropies have also found use as alternate measures of an informational content. For instance, the entropic uncertainty principle has been expressed in terms of both the Rényi [43] and the Tsallis entropies [23, 32]. Studies of generalized entropies allow to fit some properties of the standard entropy. The connection between strong subadditivity of the von Neumann entropy and the Wigner-Yanase-Dyson conjecture is a remarkable example (see [24, 41] and references therein).

The relative entropy, or Kullback-Leibler divergence [21], is frequently used as a measure of statistical distinguishability. Csiszar's  $f$ -divergence [8, 9] and Petz's quasi-entropies [25, 27] are famous generalizations of the Kullback-Leibler measure to the classical and quantum cases, respectively. In both the classical and quantum regimes, properties of the relative entropy are the subject of active research. So, the development of a standard background to generalized entropies consists an important issue. In the present paper, we examine continuity properties of the Tsallis relative entropy. The obtained bounds are expressed in terms of the trace distance between two probability distributions or density operators. In this regard, our bounds characterize a continuity property in the sense of Fannes [2].

The paper is organized as follows. In Section II, the main definitions are given. One consequence of the monotonicity of the quantum  $f$ -divergence is considered in Section III. A family of lower continuity bounds on the Tsallis relative entropy of order  $\alpha \in (0; 1)$  is derived in Section IV. In their essence, these inequalities are one-parametric extensions of the Pinsker inequality. The case of Rényi relative entropy is considered as well. In Section V, upper continuity bounds on the Tsallis relative  $\alpha$ -entropy of two probability distributions in terms of the minimal probability in its second argument are obtained. Fano type upper bounds on the conditional Tsallis entropy are derived for all  $\alpha > 0$  in Section VI. As is shown, these bounds lead to generalized inequalities of Fannes type.

## II. DEFINITIONS AND NOTATION

In the classical regime, we will consider probability distributions over the finite index set  $\Omega$  of cardinality  $N$ . The trace distance between probability distributions  $P = \{p(x)\}$  and  $Q = \{q(x)\}$  are then defined as

$$D(P, Q) := \frac{1}{2} \sum_{x \in \Omega} |p(x) - q(x)|. \quad (2.1)$$

Let  $\mathcal{L}(\mathcal{H})$  be the space of linear operators on finite-dimensional Hilbert space  $\mathcal{H}$ . We also use the notations  $\mathcal{L}_+(\mathcal{H})$  to denote the positive semidefinite operators. For any operator  $X$ , we put  $|X| \in \mathcal{L}_+(\mathcal{H})$  as a unique positive square root of  $X^*X \geq 0$ . The trace norm  $\|X\|_1 := \text{tr}|X|$  and the trace distance, defined as

$$D(X, Y) := \frac{1}{2} \|X - Y\|_1 \equiv \frac{1}{2} \text{tr}|X - Y|, \quad (2.2)$$

are widely used in both the mathematical physics and quantum information theory. Using the Ky Fan norms, the partitioned versions of the above measures can be adopted properly [29, 31]. By  $\ker(X)$  and  $\text{supp}(X)$  we denote the

kernel and the support of operator  $\mathbf{X}$ . Eigenvalues of the operator  $\mathbf{X}$  form the multi-set  $\text{spec}(\mathbf{X})$ . For two operators  $\mathbf{X}$  and  $\mathbf{Y}$  on  $\mathcal{H}$ , we define the Hilbert–Schmidt inner product by

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\text{hs}} := \text{tr}(\mathbf{X}^* \mathbf{Y}) . \quad (2.3)$$

For positive  $\alpha \neq 1$ , the Tsallis  $\alpha$ -entropy of probability distribution  $P = \{p(x)\}$  is defined by [38]

$$H_\alpha(P) := \frac{1}{1-\alpha} \left( \sum_{x \in \Omega} p(x)^\alpha - 1 \right) \equiv - \sum_{x \in \Omega} p(x)^\alpha \ln_\alpha p(x) , \quad (2.4)$$

where  $\ln_\alpha z \equiv (z^{1-\alpha} - 1)/(1-\alpha)$  is the  $\alpha$ -logarithm. The maximal value  $\ln_\alpha N$  is reached for the uniform distribution, when  $p(x) = 1/N$  for all  $x \in \Omega$ . The Shannon entropy  $H_1(P) = - \sum_x p(x) \ln p(x)$  is obtained in the limit  $\alpha \rightarrow 1$ . By  $h_\alpha(u)$  we denote the binary Tsallis  $\alpha$ -entropy, that is

$$h_\alpha(u) := H_\alpha(\{u, 1-u\}) = -u^\alpha \ln_\alpha u - (1-u)^\alpha \ln_\alpha(1-u) \quad (u \in [0; 1]) . \quad (2.5)$$

This function is concave, since its second derivative is negative. It is clear that  $h_\alpha(u) = h_\alpha(1-u)$ . Another important one-parametric generalization is the Rényi entropy (see, e.g., [9, 26] and references therein)

$$R_\alpha(P) := \frac{1}{1-\alpha} \ln \left( \sum_{x \in \Omega} p(x)^\alpha \right) . \quad (2.6)$$

The entropies (2.4) and (2.6) are connected by the equality  $(1-\alpha)R_\alpha(P) = \ln[1 + (1-\alpha)H_\alpha(P)]$ . The quantum analogs of these entropies are respectively defined as

$$H_\alpha(\rho) := \frac{1}{1-\alpha} \left( \text{tr}(\rho^\alpha) - 1 \right) , \quad (2.7)$$

$$R_\alpha(\rho) := \frac{1}{1-\alpha} \ln[\text{tr}(\rho^\alpha)] . \quad (2.8)$$

Subadditivity of the quantum Tsallis entropy (2.7) for  $\alpha > 1$  has been conjectured by Raggio [28] and later proved by Audenaert [1]. This result has been extended to some of so-called unified entropies [33]. The subadditivity property was generally believed to be true for the Wigner-Yanase entropy, until counterexamples were given [17, 37]. Meantime, if the bipartite state is pure then it is sufficient for the subadditivity. Other sufficient conditions for subadditivity of the Wigner-Yanase entropy are obtained in [7].

The standard relative entropy of  $P = \{p(x)\}$  to  $Q = \{q(x)\}$  is defined as  $H_1(P||Q) = - \sum_x p(x) \ln[q(x)/p(x)]$  [9]. For density operators  $\rho$  and  $\sigma$ , the quantum relative entropy is expressed as [26]

$$H_1(\rho||\sigma) := \text{tr}(\rho \ln \rho - \rho \ln \sigma) . \quad (2.9)$$

In the classical regime, the Tsallis relative  $\alpha$ -entropy is introduced by [4]

$$H_\alpha(P||Q) := - \sum_{x \in \Omega} p(x) \ln_\alpha \frac{q(x)}{p(x)} = \frac{1}{1-\alpha} \left( 1 - \sum_{x \in \Omega} p(x)^\alpha q(x)^{1-\alpha} \right) . \quad (2.10)$$

Basic properties of this measure are discussed in [4, 14]. The Rényi relative entropy is defined as [9]

$$R_\alpha(P||Q) := \frac{-1}{1-\alpha} \ln \left( \sum_{x \in \Omega} p(x)^\alpha q(x)^{1-\alpha} \right) . \quad (2.11)$$

It is convenient to extend the definition (2.10) to any positive-valued functions  $A$  and  $B$  on the finite set  $\Omega$ . For given set  $A = \{a(x)\}$ , we put the index subset  $\Omega_A = \{x : a(x) \neq 0\}$  and its complement  $\tilde{\Omega}_A$ . For  $\alpha > 1$ , the "Tsallis relative  $\alpha$ -entropy" of  $A = \{a(x)\}$  to  $B = \{b(x)\}$  is defined as

$$H_\alpha(A||B) := \begin{cases} \frac{1}{\alpha-1} \left( \sum_{x \in \Omega_A} a(x)^\alpha b(x)^{1-\alpha} - \sum_{x \in \Omega_A} a(x) \right) , & \tilde{\Omega}_B \subset \tilde{\Omega}_A , \\ +\infty , & \text{otherwise} . \end{cases} \quad (2.12)$$

Omitting the second entry, we obtain the definition for  $0 < \alpha < 1$ . For any positive scalar  $\lambda$ , we have

$$H_\alpha(\lambda A||\lambda B) = \lambda H_\alpha(A||B) , \quad (2.13)$$

i.e. it is a homogeneous function of degree one. For  $\alpha \in (0; 1)$  and density operators  $\rho$  and  $\sigma$ , we define the Tsallis relative entropy as

$$H_\alpha(\rho||\sigma) := \frac{1}{1-\alpha} \left( 1 - \text{tr}(\rho^\alpha \sigma^{1-\alpha}) \right). \quad (2.14)$$

For  $\alpha > 1$ , the right-hand side of (2.14) is well-defined whenever  $\ker(\sigma) \in \ker(\rho)$ . In the singular case, when the term  $\ker(\sigma) \cap \text{supp}(\rho) \neq \emptyset$  occurs, the right-hand side of (2.14) is dealt similar to the standard relative entropy (2.9). Namely, relative entropies are defined to be  $+\infty$ . Extending (2.11) to the quantum case, we define

$$R_\alpha(\rho||\sigma) := \frac{-1}{1-\alpha} \ln[\text{tr}(\rho^\alpha \sigma^{1-\alpha})]. \quad (2.15)$$

For these entropies, we have the equality

$$(\alpha - 1)R_\alpha(\rho||\sigma) = \ln[1 + (\alpha - 1)H_\alpha(\rho||\sigma)], \quad (2.16)$$

and the same relation in classical setting. For  $\alpha > 1$  and  $A, B \in \mathcal{L}_+(\mathcal{H})$ , we also introduce

$$H_\alpha(A||B) := \begin{cases} \frac{1}{\alpha-1} \left( \text{tr}(A^\alpha B^{1-\alpha}) - \text{tr}(A) \right), & \ker(B) \subset \ker(A), \\ +\infty, & \text{otherwise.} \end{cases} \quad (2.17)$$

### III. A CONSEQUENCE OF MONOTONICITY OF THE $f$ -DIVERGENCE

In this section, we prove one result which will be used for obtaining quantum bounds of Pinsker type. The Tsallis relative entropy (2.10) is closely related to the Csiszár  $f$ -divergence [8]. Let  $z \mapsto f(z)$  be a convex function on  $z \in [0; +\infty)$  with  $f(1) = 0$ . The Csiszár  $f$ -divergence of  $P = \{p(x)\}$  from  $Q = \{q(x)\}$  is defined as [8, 9]

$$S_f(P||Q) := \sum_{x \in \Omega} q(x) f\left(\frac{p(x)}{q(x)}\right). \quad (3.1)$$

Taking  $f_\alpha(z) = (\alpha - 1)^{-1} z^\alpha$  with positive  $\alpha \neq 1$  and adding corresponding constant, the formula (3.1) leads to the right-hand side of (2.10). The definition (3.1) can generally be used without the normalization condition.

In the following, we use the convention that powers of a positive semidefinite operator are only taken on its support. Namely, by  $A^{-1}$  and  $A^0$  we respectively denote the generalized inverse of  $A$  and the projection onto its support. A quantum counterpart of Csiszár's  $f$ -divergence is introduced as follows [18]. For an operator  $A \in \mathcal{L}_+(\mathcal{H})$ , let  $\Lambda_A$  and  $\Upsilon_A$  denote the left and the right multiplications by  $A$ , respectively, defined as

$$\Lambda_A : X \mapsto AX, \quad \Upsilon_A : X \mapsto XA, \quad X \in \mathcal{L}(\mathcal{H}). \quad (3.2)$$

Left and right multiplications commute with each other, namely  $\Lambda_A \Upsilon_B = \Upsilon_B \Lambda_A$  for  $A, B \in \mathcal{L}_+(\mathcal{H})$ . Let  $z \mapsto f(z)$  be a continuous function on  $z \in [0; +\infty)$ . Taking the set  $\{ab^{-1} : a \in \text{spec}(A), b \in \text{spec}(B)\}$ , we write [18]

$$f(\Lambda_A \Upsilon_{B^{-1}}) := \sum_{a \in \text{spec}(A)} \sum_{b \in \text{spec}(B)} f(ab^{-1}) \Lambda_{P_a} \Upsilon_{Q_b}, \quad (3.3)$$

where the formulas  $A = \sum_a a P_a$  and  $B = \sum_b b Q_b$  express the spectral decompositions of  $A$  and  $B$ , respectively. If  $\ker(B) \subset \ker(A)$ , then the  $f$ -divergence of  $A$  with respect to  $B$  is defined as [18]

$$S_f(A||B) := \langle B^{1/2}, f(\Lambda_A \Upsilon_{B^{-1}}) B^{1/2} \rangle_{\text{hs}}. \quad (3.4)$$

Let  $\mathbb{1}$  be the identity operator. In general case, the quantum  $f$ -divergence is defined by the formula

$$S_f(A||B) := \lim_{\varepsilon \searrow 0} S_f(A||B + \varepsilon \mathbb{1}). \quad (3.5)$$

Basic properties of the quantity (3.4) are discussed in the paper [18]. Using the function  $f_\alpha(z) = (\alpha - 1)^{-1} z^\alpha$ , for  $\ker(B) \subset \ker(A)$  we obtain

$$S_\alpha(A||B) = \frac{1}{\alpha - 1} \langle B^{1/2}, (\Lambda_A \Upsilon_{B^{-1}})^\alpha B^{1/2} \rangle_{\text{hs}} = \frac{1}{\alpha - 1} \text{tr}(A^\alpha B^{1-\alpha}). \quad (3.6)$$

Adding the term  $(1 - \alpha)^{-1} \text{tr}(\mathbf{A})$  in the right-hand side of (3.6), we have  $H_\alpha(\mathbf{A}||\mathbf{B})$  in view of (2.17). One of the most important properties of relative entropies is their monotonicity under the action of trace-preserving completely positive (TPCP) maps [39]. For a general discussion of a role of stochastic maps in quantum theory, see the paper [6]. Many fundamental results of quantum information theory are closely related to the monotonicity of the standard relative entropy [20, 24, 40]. General conditions for the monotonicity of the quantum  $f$ -divergence are obtained in [18]. If the map  $\Phi$  is TCP-map and the function  $f$  is operator convex on  $[0; +\infty)$  then

$$S_f(\Phi(\mathbf{A})||\Phi(\mathbf{B})) \leq S_f(\mathbf{A}||\mathbf{B}) . \quad (3.7)$$

Note that the inequality (3.7) has generally been established in [18] under weaker conditions on the maps. From the monotonicity (3.7) we can derive a simple upper bounds on the quantum  $f$ -divergence in terms of classical one.

**Theorem III.1.** *Let  $\mathbf{A}, \mathbf{B} \in \mathcal{L}_+(\mathcal{H})$ , and let  $\Pi_\pm$  be projectors on the eigenspaces corresponding to positive and negative eigenvalues of the difference  $(\mathbf{A} - \mathbf{B})$ . There holds*

$$S_f(\mathbf{A}||\mathbf{B}) \geq S_f(\{u'_\pm\}||\{v'_\pm\}) , \quad (3.8)$$

where  $u'_\pm = \text{tr}(\Pi_\pm \mathbf{A})$  and  $v'_\pm = \text{tr}(\Pi_\pm \mathbf{B})$ .

**Proof.** Using the Jordan decomposition of Hermitian operator

$$\mathbf{A} - \mathbf{B} = \sum_{\mu>0} \mu |\mu\rangle\langle\mu| - \sum_{\nu>0} \nu |\nu\rangle\langle\nu| , \quad (3.9)$$

we define projectors  $\Pi_+ = \sum_\mu |\mu\rangle\langle\mu|$  and  $\Pi_- = \sum_\nu |\nu\rangle\langle\nu|$ . When the difference  $(\mathbf{A} - \mathbf{B})$  has zero eigenvalues, corresponding eigenvectors should be included to the orthonormal sets  $\{|\mu\rangle\}$  and  $\{|\nu\rangle\}$  anyhow; then  $\Pi_+ + \Pi_- = \mathbf{1}$ . Consider the TCP-map  $\Psi$  defined as

$$\mathbf{A} \mapsto \Psi(\mathbf{A}) = \sum_\mu |\mu\rangle\langle\mu| \mathbf{A} |\mu\rangle\langle\mu| + \sum_\nu |\nu\rangle\langle\nu| \mathbf{A} |\nu\rangle\langle\nu| = \sum_\mu u_\mu |\mu\rangle\langle\mu| + \sum_\nu u_\nu |\nu\rangle\langle\nu| , \quad (3.10)$$

where probabilities  $u_\mu = \langle\mu|\mathbf{A}|\mu\rangle$  and  $u_\nu = \langle\nu|\mathbf{A}|\nu\rangle$ . Putting  $v_\mu = \langle\mu|\mathbf{B}|\mu\rangle$  and  $v_\nu = \langle\nu|\mathbf{B}|\nu\rangle$ , we further write

$$\Psi(\mathbf{B}) = \sum_\mu v_\mu |\mu\rangle\langle\mu| + \sum_\nu v_\nu |\nu\rangle\langle\nu| . \quad (3.11)$$

So the outputs  $\Psi(\mathbf{A})$  and  $\Psi(\mathbf{B})$  are diagonal in the same basis. Combining this fact with the inequality (3.7), we obtain

$$S_f(\mathbf{A}||\mathbf{B}) \geq S_f(\Psi(\mathbf{A})||\Psi(\mathbf{B})) = S_f(\{u_\mu, u_\nu\}||\{v_\mu, v_\nu\}) . \quad (3.12)$$

The final step is to use the monotonicity in classical regime. Let us put the 2-by- $d$  transition probability matrix

$$\mathbf{T} = \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \end{pmatrix} , \quad (3.13)$$

in which the units of the first row act on  $\mu$ -components, the units of the second row act on  $\nu$ -components. This matrix maps the ordered sets  $\{u_\mu, u_\nu\}$  and  $\{v_\mu, v_\nu\}$  to  $\{u'_+, u'_-\}$  and  $\{v'_+, v'_-\}$ , respectively. It is clear that

$$u'_+ = \sum_\mu u_\mu = \text{tr}(\Pi_+ \mathbf{A}) , \quad u'_- = \sum_\nu u_\nu = \text{tr}(\Pi_- \mathbf{A}) , \quad (3.14)$$

and analogously for  $v'_\pm$  with  $\mathbf{B}$  instead of  $\mathbf{A}$ . By the monotonicity, the right-hand side of (3.12) is larger than or equal to the right-hand side of (3.8). ■

The function  $z \mapsto z^\alpha$  is operator concave on  $\mathcal{L}_+(\mathcal{H})$  for  $0 < \alpha < 1$  (see, e.g., theorem 4.2.3 in [3]). So the function  $f_\alpha(z) = (\alpha - 1)^{-1} z^\alpha$  is operator convex for  $0 < \alpha < 1$ . Combining this with the inequality (3.8), we then obtain

$$S_\alpha(\mathbf{A}||\mathbf{B}) \geq S_\alpha(\{u'_\pm\}||\{v'_\pm\}) . \quad (3.15)$$

For density operators, we have  $u'_\pm = \text{tr}(\Pi_\pm \rho)$  and  $v'_\pm = \text{tr}(\Pi_\pm \sigma)$ . Up to a notation, the result (3.15) with density operators was presented in [30] (see theorem IV.1 therein). Writing the relation

$$\|\mathbf{A} - \mathbf{B}\|_1 = |u'_+ - v'_+| + |u'_- - v'_-| , \quad (3.16)$$

we should estimate the right-hand of (3.8) from below in terms of the distance  $\|\mathbf{A} - \mathbf{B}\|_1$ .

#### IV. PINSKER TYPE INEQUALITIES FOR $\alpha \in (0; 1)$

Studies of distinguishability measures and relations between them consist an actual issue of quantum information theory. The Pinsker inequality [8] and its quantum analog expressed as [19]

$$H_1(\rho||\sigma) \geq 2 D(\rho, \sigma)^2, \quad (4.1)$$

are well-known results of such a kind. Various lower and upper bounds on the relative entropy (2.9) were obtained in Ref. [2]. As given in terms of difference distances, these bounds characterize a continuity property in the sense of Fannes [2]. Recall that Fannes' inequality bounds from above a potential change of the von Neumann entropy in terms of trace norm distance [10]. Fannes' inequality has been extended to the Tsallis  $q$ -entropy [15, 42] and its partial sums [29]. The authors of the paper [36] proved the inequalities

$$H_\alpha(\rho||\sigma) \leq H_1(\rho||\sigma) \leq H_\beta(\rho||\sigma), \quad (4.2)$$

where  $0 \leq \alpha < 1$  and  $1 < \beta \leq 2$ . So, for  $1 < \beta \leq 2$  the relative entropy  $H_\beta(\rho||\sigma)$  is bounded from below by the right-hand side of Eq. (4.1). More detailed lower bounds on the relative entropy (2.9) are presented in Ref. [2]. By (4.2), these lower bounds are all valid for  $H_\beta(\rho||\sigma)$  with  $1 < \beta \leq 2$ .

Let  $\Pi_+$  be a projector on the eigenspace corresponding to positive eigenvalues of the difference  $(\rho - \sigma)$ . For normalized density operators, the inequality (3.15) together with the definitions (2.10) and (2.14) leads to the bound

$$H_\alpha(\rho||\sigma) \geq H_\alpha(\{u, 1-u\}||\{v, 1-v\}), \quad (4.3)$$

where we write  $u = \text{tr}(\Pi_+ \rho)$  and  $v = \text{tr}(\Pi_+ \sigma)$  for brevity. Denoting  $t = |u - v|$ , we also have  $\|\rho - \sigma\|_1 = 2t$  and  $D(\rho, \sigma) = t$ . In the paper [30], for  $u, v \in [0; 1]$  we have proved the inequality (see lemma IV.2 therein)

$$\sqrt{uv} + \sqrt{(1-u)(1-v)} \leq \sqrt{1-t^2}, \quad (4.4)$$

whence  $H_{1/2}(\{u, 1-u\}||\{v, 1-v\}) \geq 2(1 - \sqrt{1-t^2})$  and, therefore,  $H_{1/2}(\rho||\sigma) \geq D(\rho, \sigma)^2$ . We shall now estimate from below the Tsallis relative entropy (2.17) for arbitrary  $\alpha \in (0; 1)$ .

**Lemma IV.1.** *Let  $u, v \in [0; 1]$  and  $g(t) = 1 - \sqrt{1-t^2}$ . For  $\alpha \in [0; 1/2]$  and  $t = |u - v|$ , there holds*

$$u^\alpha v^{1-\alpha} + (1-u)^\alpha (1-v)^{1-\alpha} \leq 1 - 2\alpha g(t). \quad (4.5)$$

**Proof.** For fixed  $u$  and  $v$ , we define the function

$$\Phi_{uv}(\alpha) = u^\alpha v^{1-\alpha} + (1-u)^\alpha (1-v)^{1-\alpha} + 2\alpha g(t) - 1. \quad (4.6)$$

The claim (4.5) is equivalent to the inequality  $\Phi_{uv}(\alpha) \leq 0$  for all  $\alpha \in [0; 1/2]$ . First, we have  $\Phi_{uv}(0) = 0$  obviously; second,  $\Phi_{uv}(1/2) \leq 0$  in view of the relation (4.4). Third,  $\Phi_{uv}(\alpha)$  is a convex function of the parameter  $\alpha$ . Indeed, for  $u, v \neq 0, 1$  we write down

$$\frac{\partial^2 \Phi_{uv}}{\partial \alpha^2} = u^\alpha v^{1-\alpha} \left( \ln \frac{u}{v} \right)^2 + (1-u)^\alpha (1-v)^{1-\alpha} \left( \ln \frac{1-u}{1-v} \right)^2 \geq 0. \quad (4.7)$$

If a convex function is negative at the end points of some interval, it is negative in this interval everywhere. ■

Combining this with the statement of Lemma IV.1 gives a lower continuity estimate on the Tsallis relative  $\alpha$ -entropy for  $\alpha \in (0; 1)$ . We formulate for two positive operators with equal traces.

**Theorem IV.2.** *Let  $A, B \in \mathcal{L}_+(\mathcal{H})$ ,  $\text{tr}(A) = \text{tr}(B) = \theta$ ,  $D(A, B) = \tau$  and  $g(t) = 1 - \sqrt{1-t^2}$ . For all  $\alpha \in (0; 1)$ , there holds*

$$H_\alpha(A||B) \geq \varkappa_\alpha \theta g(\tau/\theta), \quad (4.8)$$

where the factor  $\varkappa_\alpha = 2\alpha(1-\alpha)^{-1}$  for  $0 < \alpha \leq 1/2$  and  $\varkappa_\alpha = 2$  for  $1/2 \leq \alpha < 1$ .

**Proof.** Using the theorem precondition and (3.16), we have  $u'_+ + u'_- = v'_+ + v'_- = \theta$  and  $\|A - B\|_1 = 2(u'_+ - v'_+)$ , whence  $\tau = u'_+ - v'_+$ . It then follows from (3.15), that

$$(1-\alpha) S_\alpha(A||B) \geq (1-\alpha) S_\alpha(\{u'_\pm\}||\{v'_\pm\}) = -\theta [u^\alpha v^{1-\alpha} + (1-u)^\alpha (1-v)^{1-\alpha}], \quad (4.9)$$

where  $u = u'_+/ \theta$ ,  $v = v'_+/ \theta$ , and  $u - v = \tau/\theta$ . Since the map (3.10) is trace-preserving,  $\text{tr}(\Psi(A)) = \theta$ . Combining this with (2.17) and (4.9) leads to

$$(1 - \alpha) H_\alpha(A||B) \geq \theta [1 - u^\alpha v^{1-\alpha} - (1 - u)^\alpha (1 - v)^{1-\alpha}] , \quad (4.10)$$

Due to the inequality (4.5), for  $0 < \alpha \leq 1/2$  the right-hand side of (4.10) is not less than  $2\alpha \theta g(\tau/\theta)$ . Hence the claim (4.8) with  $\varkappa_\alpha = 2\alpha(1 - \alpha)^{-1}$  follows. For  $1/2 \leq \alpha < 1$ , we put  $\beta = 1 - \alpha$  and also write

$$\beta H_\alpha(\rho||\sigma) \geq \theta [1 - u^{1-\beta} v^\beta - (1 - u)^{1-\beta} (1 - v)^\beta] \geq 2\beta \theta g(\tau/\theta) . \quad (4.11)$$

Hence the claim (4.8) with  $\varkappa_\alpha = 2$  follows. ■

For probability distributions, the lower continuity bound (4.8) is rewritten with the classical trace distance  $\tau = D(P, Q)$ . Expanding the function  $g(\tau/\theta)$  into power series, we obtain a family of lower bounds of the Pinsker type. Due to the binomial theorem, we actually have

$$H_\alpha(A||B) \geq \varkappa_\alpha \sum_{n=1}^{\infty} \binom{1/2}{n} (-1)^{n+1} \frac{\tau^{2n}}{\theta^{2n-1}} . \quad (4.12)$$

For all  $n$ , the coefficient  $\binom{1/2}{n} (-1)^{n+1}$  is positive. So, any partial sum of the series (4.12) provides a lower continuity bound. In particular, for normalized density operators we have a quadratic bound

$$H_\alpha(\rho||\sigma) \geq \frac{\varkappa_\alpha}{2} D(\rho, \sigma)^2 . \quad (4.13)$$

For  $1/2 \leq \alpha < 1$ , the multiplier  $\varkappa_\alpha/2 = 1$ , whereas for  $\alpha = 1$  we have the quadratic bound (4.1) with the multiplier two. This is evidence for that the series (4.12) does not provide an expansion with the best constants at powers of the trace distance. For the standard relative entropy  $H_1(P||Q)$ , these constants have been the subject of long-time research (for details, see [12] and references therein). It would be interesting to develop methods for finding the best constants in Pinsker type inequalities for the Tsallis relative  $\alpha$ -entropy. For instance, we could try to extend those ways that are known for the case  $\alpha = 1$ . Nevertheless, the results (4.8) and (4.12) resolve the question in a general sense, since they provide non-trivial lower continuity bounds on the Tsallis relative  $\alpha$ -entropy. For  $\alpha \in (0; 1)$ , we also combine (2.16) with (4.8) and hence obtain

$$R_\alpha(\rho||\sigma) \geq \frac{1}{\alpha - 1} \ln[1 - (1 - \alpha)\varkappa_\alpha g(\tau)] \geq \varkappa_\alpha g(\tau) , \quad (4.14)$$

where  $\tau = D(\rho, \sigma)$ . Indeed, the function  $(\alpha - 1)^{-1} \ln[1 - (1 - \alpha)x]$  increases with  $x \in [0; \frac{1}{1-\alpha})$  and  $-\ln(1 - \xi) \geq \xi$  for  $\xi \in [0; 1)$ . The inequality (4.14) can be regarded as the bound of Pinsker type on the Rényi relative entropy for  $\alpha \in (0; 1)$ . In particular, we have a version of (4.13) with  $R_\alpha(\rho||\sigma)$  instead of  $H_\alpha(\rho||\sigma)$ . Thus, we have obtained a family of lower continuity bounds in terms of the trace distance on both the relative entropies (2.14) and (2.15).

## V. UPPER CONTINUITY BOUNDS FOR $\alpha > 1$

One of basic features of the standard relative entropy is its unboundedness. The relative  $\alpha$ -entropy enjoys the same for  $\alpha > 1$ . So we may ask a behaviour of the functional  $H_\alpha(P||Q)$  as the minimal probability in  $Q$  goes to zero. Of course, in the quantum case this question is more difficult due to the non-commutativity. For the standard relative entropy, such an upper bound was presented in [5], and more bounds were given in [2]. For the quantum relative  $\alpha$ -entropy of order  $\alpha > 1$ , upper continuity bounds in terms of the minimal eigenvalue of its second entry were obtained in [34]. It turns out that in the commutative case these bounds can be sharpened significantly. Our derivation will mainly be based on the joint convexity. Namely, for each positive  $\alpha \neq 1$  the quantity (2.12) enjoys

$$H_\alpha(\theta A^{(1)} + (1 - \theta)A^{(2)} || \theta B^{(1)} + (1 - \theta)B^{(2)}) \leq \theta H_\alpha(A^{(1)}||B^{(1)}) + (1 - \theta)H_\alpha(A^{(2)}||B^{(2)}) \quad (5.1)$$

for all  $0 \leq \theta \leq 1$ . This relation can be proved by means of so-called "generalized log-sum inequality" (see formula (16) in [4]). The properties (2.13) and (5.1) lead to the following statement.

**Lemma V.1.** *Let  $A, B, C$  be three sets of positive numbers, and  $\mathbf{A}, \mathbf{B}$ , and  $\mathbf{C}$  three positive operators. There holds*

$$H_\alpha(A + C||B + C) \leq H_\alpha(A||B) \quad (0 \leq \alpha < \infty) , \quad (5.2)$$

$$H_\alpha(\mathbf{A} + \mathbf{C}||\mathbf{B} + \mathbf{C}) \leq H_\alpha(\mathbf{A}||\mathbf{B}) \quad (0 \leq \alpha \leq 2) . \quad (5.3)$$

**Proof.** Using (2.13) and (5.1), we merely write

$$H_\alpha(A + C||B + C) = 2 H_\alpha((A + C)/2 || (B + C)/2) \leq H_\alpha(A||B) + H_\alpha(C||C) = H_\alpha(A||B) \quad (5.4)$$

in view of obvious  $H_\alpha(C||C) = 0$ . The quantum relative entropy (2.17) also enjoys both the homogeneity of degree one and the joint convexity, but the latter only for  $0 \leq \alpha \leq 2$  (see, e.g., the review [20]). Rewriting the above arguments with the quantum relative  $\alpha$ -entropy instead of the classical one, we have arrived at the claim (5.3). ■

For the standard relative entropy (2.9), the relation (5.3) was proved in [2]. The inequality (5.2) can be utilized for obtaining an upper bound on  $H_\alpha(P||Q)$  in terms of the trace distance  $D(P, Q)$  and the minimal probability

$$q_0 := \min\{q_j : j \in \Omega_P\} . \quad (5.5)$$

Here we apply that any sum in  $H_\alpha(P||Q)$  is effectively restricted to the index subset  $\Omega_P$ . Defining the set  $\Delta = P - Q$  with elements  $\delta_j = p_j - q_j$ , we put another set  $\bar{Q}$  with positive elements

$$\bar{q}_j := \max\{q_0, -\delta_j\} . \quad (5.6)$$

Writing  $Q = \bar{Q} + (Q - \bar{Q})$  and using the property (5.2) with  $C = Q - \bar{Q}$ , we obtain

$$H_\alpha(P||Q) = H_\alpha(\Delta + \bar{Q} + (Q - \bar{Q}) || \bar{Q} + (Q - \bar{Q})) \leq H_\alpha(\Delta + \bar{Q} || \bar{Q}) . \quad (5.7)$$

Use of (5.2) is correct here due to positivity of both  $\Delta + \bar{Q}$  and  $Q - \bar{Q}$ : we clearly have  $\delta_j + \max\{q_0, -\delta_j\} \geq 0$  and  $q_j - \max\{q_0, q_j - p_j\} \geq 0$ . The maximization of the right-hand side of (5.7) is under the conditions  $\sum_j \delta_j = 0$  and  $\sum_j |\delta_j| = 2D(P, Q)$ . We separately consider the two cases,  $D(P, Q) \leq q_0$  and  $q_0 < D(P, Q) \leq 1 - q_0$ .

**Theorem V.2.** Let  $q_0$  be defined by (5.5),  $\tilde{\Omega}_Q \subset \tilde{\Omega}_P$  and  $\tau = D(P, Q)$ . For  $\alpha > 1$ , the Tsallis relative  $\alpha$ -entropy is bounded from above as

$$H_\alpha(P||Q) \leq \frac{1}{\alpha - 1} \left( (q_0 + \tau)^\alpha q_0^{1-\alpha} + (q_0 - \tau)^\alpha q_0^{1-\alpha} - 2q_0 \right) \quad (\tau \leq q_0) , \quad (5.8)$$

$$H_\alpha(P||Q) \leq \frac{1}{\alpha - 1} \left( (q_0 + \tau)^\alpha q_0^{1-\alpha} - (q_0 + \tau) \right) \quad (q_0 < \tau \leq 1 - q_0) . \quad (5.9)$$

**Proof.** In the case  $\tau \leq q_0$ , we have  $\max(q_0, -\delta_j) = q_0$  for all  $j$ . By  $x_i$  and  $(-y_j)$  we respectively denote positive and negative elements of the set  $\Delta = \{\delta_j\}$ . The conditions  $\sum_j \delta_j = 0$  and  $\sum_j |\delta_j| = 2\tau$  are rewritten as

$$\sum_{i \in \omega_x} x_i = \sum_{j \in \omega_y} y_j = \tau , \quad (5.10)$$

where  $\omega_x$  and  $\omega_y$  are corresponding subsets of the set  $\Omega_P$  of cardinality  $n$ . The right-hand side of (5.7) is represented as the function

$$F(x_i, y_j) = \frac{1}{\alpha - 1} \left( \sum_{i \in \omega_x} (q_0 + x_i)^\alpha q_0^{1-\alpha} + \sum_{j \in \omega_y} (q_0 - y_j)^\alpha q_0^{1-\alpha} - nq_0 \right) . \quad (5.11)$$

This function should be maximized under the conditions  $0 \leq x_i$ ,  $0 \leq y_j$  and (5.10), which define a simplex. Recall that the global maximum of a convex function relative to a convex set is reached at one of the extreme points of that set [35]. Hence the maximum of  $F(x_i, y_j)$  is equal to the right-hand side of (5.8) and reached, when one of the  $x_i$ 's and one of the  $y_j$ 's are equal to  $\tau$  and other are all zero.

In the case  $q_0 < \tau$ , we add negative elements  $(-z_k)$  such that  $q_0 < z_k$  and  $\sum_{j \in \omega_y} y_j + \sum_{k \in \omega_z} z_k = \tau$ . Before maximization, we further simplify the right-hand side of (5.7). Putting the sets  $\Delta_x = \{x_i\}$ ,  $\Delta_y = \{y_j\}$  and  $\Delta_z = \{z_k\}$ , we write

$$\Delta = \Delta_x - \Delta_y - \Delta_z , \quad \bar{Q} = q_0(I_x + I_y) + \Delta_z , \quad \Delta + \bar{Q} = q_0(I_x + I_y) + \Delta_x - \Delta_y . \quad (5.12)$$

Here  $I_x$  denote the indicator of the set  $\omega_x$  taking the value one for  $j \in \omega_x$  and zero for  $j \notin \omega_x$ . Note that if the set  $\Omega_Z$  does not intersect with both the  $\Omega_A$  and  $\Omega_B$  then  $H_\alpha(A||B + Z) = H_\alpha(A||B)$ . Using this fact twice and the inequality (5.2) again with positive  $C = q_0 I_y - \Delta_y$ , we rewrite the right-hand side of (5.7) in a form

$$\begin{aligned} H_\alpha(q_0(I_x + I_y) + \Delta_x - \Delta_y || q_0(I_x + I_y) + \Delta_z) &= H_\alpha(q_0 I_x + \Delta_x + C || q_0 I_x + \Delta_y + C) \\ &\leq H_\alpha(q_0 I_x + \Delta_x || q_0 I_x + \Delta_y) = H_\alpha(q_0 I_x + \Delta_x || q_0 I_x) . \end{aligned} \quad (5.13)$$

The latter is expressed as  $G(x_i) = (\alpha - 1)^{-1} (\sum_{i \in \omega_x} (q_0 + x_i)^\alpha q_0^{1-\alpha} - (n_x q_0 + \tau))$ , where  $n_x$  is cardinality of the  $\omega_x$ . Under the conditions  $0 \leq x_i$  and  $\sum_i x_i = \tau$ , the maximum of  $G(x_i)$  is equal to the right-hand side of (5.9) and reached, when one of the  $x_i$ 's is  $\tau$  and other are all zero. ■

The upper bounds (5.8) and (5.9) have a behavior  $q_0^{1-\alpha}$  with respect to the minimal probability  $q_0$ . For the quantum relative entropy  $H_\alpha(\rho||\sigma)$ , upper continuity bounds with a similar dependence on the minimal eigenvalue of  $\sigma$  were obtained in our previous paper [34]. The bounds (5.8) and (5.9) are stronger, but their proof is quite restricted to the commutative case. The principal point is that positivity of diagonal elements of a matrix do not imply positivity of matrix itself (except for the case of diagonal matrices). So, the proof of Theorem V.2 is purely classical in character. On the other hand, weaker bounds in the paper [34] have been proved just for the quantum case. Note that the inequalities (5.8) and (5.9) can be rewritten in terms of  $\alpha$ -logarithm as

$$H_\alpha(P||Q) \leq -(q_0 + \tau_0) \ln_\alpha \left( \frac{q_0}{q_0 + \tau} \right) - (q_0 - \tau) \ln_\alpha \left( \frac{q_0}{q_0 - \tau} \right) \quad (\tau \leq q_0) , \quad (5.14)$$

$$H_\alpha(P||Q) \leq -(q_0 + \tau) \ln_\alpha \left( \frac{q_0}{q_0 + \tau} \right) \quad (b_0 < \tau \leq 1 - q_0) . \quad (5.15)$$

Using (2.16) in the classical regime, we see that the bounds (5.14) and (5.15) remain valid with  $R_\alpha(P||Q)$  instead of  $H_\alpha(P||Q)$ . This claim follows from the points that  $\ln\{1 + (\alpha - 1)x\}$  increases with positive  $x$  and  $\ln(1 + \xi) \leq \xi$  for  $\xi \geq 0$ . The relations (5.14) and (5.15) are  $\alpha$ -parametric extensions of the upper bounds obtained in Ref. [2] for the standard relative entropy. In effect, the upper continuity bounds (5.8) and (5.9) could be used for classical systems which are well described within non-extensive thermostatics.

## VI. NOTES ON FANO AND FANNES INEQUALITIES

In this section, we will obtain upper bounds on the conditional Tsallis  $\alpha$ -entropy for all  $\alpha > 0$ . It is convenient to change the notation slightly as follows. Let  $X$  and  $Y$  be discrete random variables with probabilities  $\{p_X(x)\}$  and  $\{p_Y(y)\}$ , each supported on the  $N$ -point set  $\Omega$ . By  $p_{XY}(x, y)$  and  $p_{X|Y}(x|y)$  we respectively denote the joint and conditional probabilities. The joint  $\alpha$ -entropy and the conditional  $\alpha$ -entropy are respectively defined as

$$H_\alpha(X, Y) := \frac{1}{1 - \alpha} \left( \sum_{x, y} p_{XY}(x, y)^\alpha - 1 \right) , \quad (6.1)$$

$$H_\alpha(X|Y) := \sum_y p_Y(y)^\alpha H_\alpha(X|y) , \quad (6.2)$$

where  $H_\alpha(X|y) = (1 - \alpha)^{-1} (\sum_x p_{X|Y}(x|y)^\alpha - 1)$ . Rewriting  $H_\alpha(X|y) = -\sum_x p_{X|Y}(x|y)^\alpha \ln_\alpha p_{X|Y}(x|y)$ , we further obtain

$$H_\alpha(X|Y) = -\sum_{x, y} p_{XY}(x, y)^\alpha \ln_\alpha p_{X|Y}(x|y) , \quad (6.3)$$

due to the relation  $p_Y(y) p_{X|Y}(x|y) = p_{XY}(x, y)$ . We will follow the original scheme of derivation (see the classical text [11], section 6.2). The probability of error is expressed as

$$P_e = \sum_y p_Y(y) q(e|y) , \quad q(e|y) = 1 - p_{X|Y}(y|y) = \sum_{x \neq y} p_{X|Y}(x|y) . \quad (6.4)$$

**Lemma VI.1.** *For all  $\alpha \in (0; \infty)$ , there holds*

$$H_\alpha(X|Y) \leq \sum_y p_Y(y)^\alpha h_\alpha(q(e|y)) + \ln_\alpha(N - 1) \sum_y p_Y(y)^\alpha q(e|y)^\alpha . \quad (6.5)$$

**Proof.** Using the expression for  $p(e|y)$  and the definition (2.5), we write

$$\begin{aligned} H_\alpha(X|y) &= -p_{X|Y}(y|y)^\alpha \ln_\alpha p_{X|Y}(y|y) - \sum_{x \neq y} p_{X|Y}(x|y)^\alpha \ln_\alpha p_{X|Y}(x|y) \\ &= h_\alpha(q(e|y)) + q(e|y)^\alpha \ln_\alpha q(e|y) - \sum_{x \neq y} p_{X|Y}(x|y)^\alpha \ln_\alpha p_{X|Y}(x|y) . \end{aligned} \quad (6.6)$$



Due to  $q(e|y) = \sum_{x \neq y} p_{X|Y}(x|y)$  and the properties of  $\alpha$ -logarithm, the second and third terms in the right-hand side of (6.6) are combined as

$$\begin{aligned} & - \sum_{x \neq y} p_{X|Y}(x|y)^\alpha \ln_\alpha p_{X|Y}(x|y) + q(e|y) q(e|y)^{\alpha-1} \ln_\alpha q(e|y) \\ & = - \sum_{x \neq y} p_{X|Y}(x|y)^\alpha \left( \ln_\alpha p_{X|Y}(x|y) - p_{X|Y}(x|y)^{1-\alpha} q(e|y)^{\alpha-1} \ln_\alpha q(e|y) \right) \\ & = - \sum_{x \neq y} p_{X|Y}(x|y)^\alpha \left( \ln_\alpha p_{X|Y}(x|y) + p_{X|Y}(x|y)^{1-\alpha} \ln_\alpha \frac{1}{q(e|y)} \right) \end{aligned} \quad (6.7)$$

$$= -q(e|y)^\alpha \sum_{x \neq y} \frac{p_{X|Y}(x|y)^\alpha}{q(e|y)^\alpha} \ln_\alpha \frac{p_{X|Y}(x|y)}{q(e|y)} \leq q(e|y)^\alpha \ln_\alpha(N-1) . \quad (6.8)$$

Here we used the identities  $\ln_\alpha(1/\xi) = -\xi^{\alpha-1} \ln_\alpha \xi$  (right before (6.7)) and  $\ln_\alpha(\xi z) = \ln_\alpha \xi + \xi^{1-\alpha} \ln_\alpha z$  (right before (6.8)). Substituting (6.8) in (6.6) and further in (6.2), we obtain (6.5). ■

**Theorem VI.2.** *Let random variables  $X$  and  $Y$  take values on the same finite set with cardinality  $N$ . For given value of the error probability  $P_e$ , the conditional entropy  $H_\alpha(X|Y)$  is bounded from above as*

$$H_\alpha(X|Y) \leq \frac{P_e^\alpha - \alpha P_e}{1 - \alpha} + P_e^\alpha \ln_\alpha[N(N-1)] \quad (0 < \alpha < 1) , \quad (6.9)$$

$$H_\alpha(X|Y) \leq h_\alpha(P_e) + P_e^\alpha \ln_\alpha(N-1) \quad (1 < \alpha < \infty) . \quad (6.10)$$

**Proof.** For  $\alpha \in (0; 1)$ , we use the expression  $h_\alpha(u) = (1 - \alpha)^{-1}[u^\alpha + (1 - u)^\alpha - 1] \leq (1 - \alpha)^{-1}(u^\alpha - \alpha u)$ , which follows from (2.5) and the inequality

$$1 - (1 - u)^\alpha = \int_0^u \alpha(1 - t)^{\alpha-1} dt \geq \int_0^u \alpha dt = \alpha u . \quad (6.11)$$

By these relations and  $\xi_y = p_Y(y) q(e|y)$ , the first sum in the right-hand side of (6.5) is no larger than

$$\frac{1}{1 - \alpha} \sum_y p_Y(y)^\alpha [q(e|y)^\alpha - \alpha q(e|y)] \leq \frac{1}{1 - \alpha} \left( \sum_y \xi_y^\alpha - \alpha P_e \right) , \quad (6.12)$$

in view of  $\sum_y p_Y(y)^\alpha q(e|y) \geq \sum_y p_Y(y) q(e|y) = P_e$ . Using the Hölder inequality, we also obtain

$$\max \left\{ \sum_{y=1}^N \xi_y^\alpha : 0 \leq \xi_y \leq 1, \sum_{y=1}^N \xi_y = P_e \right\} = N^{1-\alpha} P_e^\alpha , \quad (6.13)$$

which is reached for  $\xi_y = P_e/N$ . So the term  $(1 - \alpha)^{-1} (N^{1-\alpha} P_e^\alpha - \alpha P_e)$  is an upper bound for the right-hand side of (6.12). Combining this with the product of  $\ln_\alpha(N-1)$  and (6.13) finally gives (6.9).

Using  $p_Y(y)^\alpha \leq p_Y(y)$  for  $\alpha > 1$  and Jensen's inequality for the concave function (2.5), we have

$$\sum_y p_Y(y)^\alpha h_\alpha(q(e|y)) \leq \sum_y p_Y(y) h_\alpha(q(e|y)) \leq h_\alpha \left( \sum_y p_Y(y) q(e|y) \right) = h_\alpha(P_e) . \quad (6.14)$$

For  $\alpha > 1$ , there holds  $\sum_y p_Y(y)^\alpha q(e|y)^\alpha \leq \left( \sum_y p_Y(y) q(e|y) \right)^\alpha = P_e^\alpha$  (this can be rewritten as  $\|\cdot\|_\alpha \leq \|\cdot\|_1$  in terms of the vector norms). By these two points, the inequality (6.5) leads to (6.10). ■

For  $\alpha > 1$ , the inequality (6.10) with  $P_e$  instead of  $P_e^\alpha$  was presented in [13]. So we obtain an improvement of the known result. The inequality (6.9) for  $0 < \alpha < 1$  is a new bound. By construction, the bound (6.9) is not sharp. Nevertheless, this inequality is sufficiently exact for small values of  $P_e$ . In any event, both the bounds of Theorem VI.2 show that  $P_e \rightarrow 0$  implies  $H_\alpha(X|Y) \rightarrow 0$ . On the other hand, if  $H_\alpha(X|Y)$  is large then the probability of making an error in inference must be large as well. In this regard, the essence of our inequalities concurs with a typical use of the standard Fano inequality in terms of the Shannon entropies.

Uniform continuity is an important property of the von Neumann entropy. The first result in this issue was given by Fannes [10]. The Tsallis entropy itself [15, 42] and its partial sums [29] also enjoy this property. Using the classical Fano inequality, Fannes' bound has been sharpened (see theorem 3.8 and its proof of Csiszár in [26]). We shall now show that the Fano type inequalities (6.9) and (6.10) lead to the Fannes inequality in terms of Tsallis entropies. Using properties of the  $\alpha$ -logarithm, the joint entropy (6.1) can be re-expressed as [13]

$$H_\alpha(X, Y) = H_\alpha(X) + H_\alpha(Y|X) = H_\alpha(Y) + H_\alpha(X|Y) . \quad (6.15)$$

Hence, in view of  $H_\alpha(Y|X) \geq 0$ , the difference  $H_\alpha(X) - H_\alpha(Y) \leq H_\alpha(X|Y)$  is bounded from above by the right-hand side of (6.9) for  $\alpha \in (0; 1)$  and by the right-hand side of (6.10) for  $\alpha \in (1; +\infty)$ . For given probability distributions  $\{p_X(x)\}$  and  $\{p_Y(y)\}$ , the joint probability mass function  $p_{XY}(x, y)$  can be built in such a way that  $P_e = D(X, Y) = (1/2) \sum_x |p_X(x) - p_Y(x)|$ , – this follows from the coupling inequality (see, e.g., the book [22]). Setting  $\{p_X(x)\} = \text{spec}(\rho)$  and  $\{p_Y(y)\} = \text{spec}(\sigma)$ , we then have  $D(X, Y) \leq D(\rho, \sigma)$  (see, e.g., lemma 11.1 in [26]). Assuming  $N \geq 2$ , the right-hand side of (6.9) increases with  $P_e$  for all  $0 \leq P_e \leq 1$ , the right-hand side of (6.10) increases with  $P_e$  for all  $0 \leq P_e \leq N/(N+1)$ . Replacing  $D(X, Y)$  with larger  $D(\rho, \sigma)$ , we get the following result.

**Theorem VI.3.** *Let  $d$  be dimensionality of the Hilbert space and  $\tau = D(\rho, \sigma)$ ; then*

$$|H_\alpha(\rho) - H_\alpha(\sigma)| \leq \frac{\tau^\alpha - \alpha\tau}{1 - \alpha} + \tau^\alpha \ln_\alpha[d(d-1)] \quad (0 < \alpha < 1, \quad 0 \leq \tau \leq 1), \quad (6.16)$$

$$|H_\alpha(\rho) - H_\alpha(\sigma)| \leq h_\alpha(\tau) + \tau^\alpha \ln_\alpha(d-1) \quad \left(1 < \alpha, \quad 0 \leq \tau \leq \frac{d}{d+1}\right). \quad (6.17)$$

The inequality (6.17), when  $\alpha > 1$ , is just the uniform estimate derived by a direct method in [42]. This inequality is the best bound known for the Tsallis entropies in the parameter range  $\alpha > 1$ . In the limit  $\alpha \rightarrow 1$ , the inequality (6.17) reproduces the statement of theorem 3.8 in [26]. For  $\alpha \in (0; 1)$ , there exists another inequality

$$|H_\alpha(\rho) - H_\alpha(\sigma)| \leq \frac{(2\tau)^\alpha - 2\tau}{1 - \alpha} + (2\tau)^\alpha \ln_\alpha d, \quad (6.18)$$

provided that  $\|\rho - \sigma\|_1 = 2\tau \leq \alpha^{1/(1-\alpha)}$ . The bound (6.18) was actually proved in the paper [15] for all  $\alpha \in [0; 2]$ , but the bound (6.17) is better for  $\alpha \geq 1$ . Comparing our bound (6.16) with (6.18), we see the following. In general, the bound (6.16) is weaker but covers all acceptable values  $\tau \in [0; 1]$  of the trace distance. The scope of (6.18) is restricted to the range  $0 \leq 2\tau \leq \alpha^{1/(1-\alpha)}$ . In low dimensions, however, the bound (6.16) can be better than the bound (6.18). Say, for  $d = 2$  and  $\alpha = 1/2$  the bound (6.18) holds for  $0 \leq \tau \leq 1/8$ . In this range, the right-hand side of (6.18) is larger than the right-hand side of (6.16). Moreover, for sufficiently small  $\tau$  the difference between the two bounds is up to 40 %. Thus, the bound (6.16) has some practical interest, at least in the primary qubit case.

- 
- [1] Audenaert, K.: Subadditivity of  $q$ -entropies for  $q > 1$ , J. Math. Phys. **48**, 083507 (2007)
  - [2] Audenaert, K.M.R., Eisert, J.: Continuity bounds on the quantum relative entropy. J. Math. Phys. **46**, 102104 (2005)
  - [3] Bhatia, R.: Positive Definite Matrices. Princeton University Press, Princeton (2007)
  - [4] Borland, L., Plastino, A.R., Tsallis, C.: Information gain within nonextensive thermostatics. J. Math. Phys. **39**, 6490–6501 (1998); **40**, 2196(E) (1999)
  - [5] Bratteli, O., Robinson, D.W.: Operator Algebras and Quantum Statistical Mechanics 2. Equilibrium States. Models in Quantum Statistical Mechanics. Springer, Berlin (2002)
  - [6] Busch, P: Stochastic isometries in quantum mechanics. Math. Phys. Anal. Geom. **2**, 83–106 (1999)
  - [7] Cai, L., Hansen, F.: Metric-adjusted skew information: convexity and restricted forms of superadditivity. Lett. Math. Phys. **93**, 1–13 (2010)
  - [8] Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hungar. **2**, 299–318 (1967)
  - [9] Csiszár, I.: Axiomatic characterizations of information measures. Entropy **10**, 261–273 (2008)
  - [10] Fannes, M.: A continuity property of entropy density for spin lattice systems. Commun. Math. Phys. **31**, 291–294 (1973)
  - [11] Fano, R.M.: Transmission of Information: a Statistical Theory of Communications. MIT Press and John Wiley & Sons, New York (1961)
  - [12] Fedotov, A., Harremoës, P., Topsøe, F.: Refinements of Pinsker inequality. IEEE Trans. Inf. Theory **49**, 1491–1498 (2003)
  - [13] Furuichi, S.: Information theoretical properties of Tsallis entropies. J. Math. Phys. **47**, 023302 (2006)
  - [14] Furuichi, S., Yanagi, K., Kuriyama, K.: Fundamental properties of Tsallis relative entropy. J. Math. Phys. **45**, 4868–4877 (2004)
  - [15] Furuichi, S., Yanagi, K., Kuriyama, K.: A generalized Fannes’ inequality. J. Inequal. Pure Appl. Math. **8**(1), 5 (2007)
  - [16] Gell-Mann, M., Tsallis, C., ed.: Nonextensive Entropy – Interdisciplinary Applications. Oxford University Press, Oxford (2004)
  - [17] Hansen, F.: The Wigner–Yanase entropy is not subadditive. J. Stat. Phys. **126**, 643–648 (2007)
  - [18] F. Hiai, M. Mosonyi, D. Petz, C. Bény.: Quantum  $f$ -divergence and error correction. E-print arXiv:1008.2529v5 [math-ph]
  - [19] Hiai, F., Ohya, M., Tsukada, M.: Sufficiency, KMS condition and relative entropy in von Neumann algebras. Pacific J. Math. **96**, 99–109 (1981)
  - [20] Jenčová, A., Ruskai, M.B.: A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality. Rev. Math. Phys. **22**, 1099–1121 (2010)

- [21] Kullback, S., Leibler, R.A.: On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951)
- [22] Lindvall, T.: *Lectures on the Coupling Method*. John Wiley & Sons, New York (1992)
- [23] Majerník, V., Majerníková, E.: The determination of bounds of the  $\beta$ -entropic sum of two noncommuting observables. *Rep. Math. Phys.* **47**, 381–392 (2001)
- [24] Nielsen, M.A., Chuang, I.L.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2000)
- [25] Petz, D.: Quasi-entropies for finite quantum systems. *Rep. Math. Phys.* **21**, 57–65 (1986)
- [26] Petz, D.: *Quantum Information Theory and Quantum Statistics*. Springer, Berlin (2008)
- [27] Petz, D.: From  $f$ -divergence to quantum quasi-entropies and their use. *Entropy* **12**, 304–325 (2010)
- [28] Raggio, G.A.: Properties of  $q$ -entropies, *J. Math. Phys.* **36**, 4785–4791 (1995)
- [29] Rastegin, A.E.: Continuity and stability of partial entropic sums. *Lett. Math. Phys.* **94**, 229–242 (2010)
- [30] Rastegin, A.E.: Fano type quantum inequalities in terms of  $q$ -entropies. *E-print* arXiv:1010.1811 [quant-ph]
- [31] Rastegin, A.E.: Bounds on Shannon distinguishability in terms of partitioned measures. *Quantum Inf. Process.* **10**, 123–138 (2011)
- [32] Rastegin, A.E.: Entropic uncertainty relations for extremal unravelings of super-operators. *J. Phys. A: Math. Theor.* **44**, 095303 (2011)
- [33] Rastegin, A.E.: Some general properties of unified entropies. *J. Stat. Phys.* **143**, 1120–1135 (2011)
- [34] Rastegin, A.E.: Upper continuity bounds on the relative  $q$ -entropy for  $q > 1$ . *J. Math. Phys.* **52**, 062203 (2011)
- [35] Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton (1970)
- [36] Ruskai, M.B., Stillinger, F.M.: Convexity inequalities for estimating free energy and relative entropy. *J. Phys. A: Math. Gen.* **23**, 2421–2437 (1990)
- [37] Seiringer, R.: On the failure of subadditivity of the Wigner–Yanase entropy. *Lett. Math. Phys.* **80**, 285–288 (2007)
- [38] Tsallis, C.: Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **32**, 479–487 (1988)
- [39] Uhlmann, A.: Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory. *Commun. Math. Phys.* **54**, 21–32 (1977)
- [40] Vedral, V.: The role of relative entropy in quantum information theory. *Rev. Mod. Phys.* **74**, 197–234 (2002)
- [41] Wehrl, A.: General properties of entropy. *Rev. Mod. Phys.* **50**, 221–260 (1978)
- [42] Zhang, Z.: Uniform estimates on the Tsallis entropies. *Lett. Math. Phys.* **80**, 171–181 (2007)
- [43] Zozor, S., Portesi, M., Vignat, C.: Some extensions of the uncertainty principle. *Physica A* **387**, 4800–4808 (2008)